Sociodemographic Factors and the Variation in Syphilis Rates among US Counties, 1984 through 1993: An Ecological Analysis

Peter H. Kilmarx, MD, Akbar A. Zaidi, MS, James C. Thomas, PhD, MPH, Allyn K. Nakashima, MD, Michael E. St. Louis, MD, Melinda L. Flock, MSPH, and Thomas A. Peterman, MD, MSc

ABSTRACT

Objectives. Syphilis in the United States is focally distributed, with high incidence rates in the South and in metropolitan areas nationwide. In this study an ecological analysis, using the county as the unit of analysis, was performed to generate hypotheses about community-level determinants of syphilis rates.

Methods. Bivariate rank correlations and multivariate, backward stepwise elimination linear regressions were performed. Mean annual incidence of primary- and secondary-stage syphilis in a county was the dependent variable, and county sociodemographic characteristics (from census data) were the independent variables.

Results. In the multivariate regression model, sociodemographic characteristics accounted for 71% of the variation in syphilis rates among counties. With other factors accounted for, the most highly correlated characteristics were percentage non-Hispanic Black population, county location in the South, percentage of the population that was urban, percentage Hispanic population, and percentage of births to women younger than 20 years.

Conclusions. Most of the variation in syphilis rates among counties is accounted for by sociodemographic characteristics. Identification and remediation of modifiable health determinants for which these factors are markers are needed to improve the health status of these populations. (*Am J Public Health.* 1997;87:1937–1943)

Introduction

The most recent syphilis epidemic in the United States peaked in 1990, with 20.3 cases of primary- and secondary-stage syphilis per 100 000 persons, the highest rate since the 1940s and more than 10 times as high as rates in other developed countries.[1,2] Research has shown syphilis to be related to demographic factors such as age, marital status, urban residence, income, education, and race/ethnicity,[3] and to behavioral factors such as numbers of sex partners, homosexual activity among men,[4] and, particularly in this last epidemic, the use of crack cocaine and the exchange of sex for drugs.[5,6]

Syphilis is focally distributed, with higher rates in the southeastern United States and in metropolitan areas nationwide while many other areas have few or no reported cases. Previous studies have used the individual as the unit of analysis and usually have not accounted for group-level factors such as population density or the activity of disease control programs. Such studies cannot address the geographic distribution of syphilis. Also, individual-level analysis may be less appropriate for transmissible infectious disease epidemiology in which infection risk depends on the prevalence of disease in a community.[7–9] Ecological analysis, which uses groups as the unit of analysis, is more appropriate for identifying community-level determinants of syphilis rates and for studying the focal distribution of syphilis. To describe sociodemographic correlates of syphilis infection rates and to generate hypotheses about community-level determinants of syphilis rates in the United States, we performed an ecological study of syphilis incidence, using the county as the unit of analysis.

Methods

Data on syphilis cases came from the national syphilis case surveillance system. Reports of suspected cases of syphilis and reactive syphilis serology reports, depending upon age of the patient and titer, are investigated by local health departments.[10] Cases are reported to state health departments, which send monthly reports to the Centers for Disease Control and Prevention (CDC) in Atlanta, Ga. Reports include aggregate totals of primary- and secondary-stage syphilis cases by county of patient's residence. Since trends in primary- and secondary-stage syphilis may be less subject than trends in latent syphilis to artifactual fluctuations caused by variation in intensity of case finding, the analyses in this study were confined to primary- and secondary-stage syphilis. Cases in Alaska are not reported at the county level and were excluded from these analyses.

County syphilis rates were calculated with the US Bureau of the Census population count (1990), intercensal estimates (1984 through 1989), and postcensal estimates (1991 through 1992) used as the population denominators. The 1992 estimate was used to calculate rates for 1993.

Peter H. Kilmarx, Akbar A. Zaidi, Allyn K. Nakashima, Michael E. St. Louis, Melinda L. Flock, and Thomas A. Peterman are with the Division of STD Prevention, Centers for Disease Control and Prevention (CDC), Atlanta, Ga. Peter H. Kilmarx is also with the CDC's Epidemic Intelligence Service. James C. Thomas is with the Department of Epidemiology, University of North Carolina School of Public Health, Chapel Hill.

Requests for reprints should be sent to Information Dissemination, Communications Office, NCHSTP, MS E-06, CDC, 1600 Clifton Rd, Atlanta, GA 30333.

This paper was accepted April 10, 1997.

Since rates are unstable in counties with small populations, the 26 US counties with 1990 populations of fewer than 1000 persons were excluded from these analyses. Because the epidemic peaked at different times in different regions in the United States, and to account for endemic as well as epidemic rates, the county 10-year mean annual primary- and secondary-stage syphilis incidence was used as the dependent outcome variable.

Data from the US Bureau of the Census were used as the independent variables in these analyses.[11,12] Available data elements were reviewed and county sociodemographic characteristics that were plausibly linked to syphilis rates were examined. Twenty county sociodemographic characteristics were analyzed as continuous variables: percentage of the population aged 14 years or younger (1990); percentage of the population aged 15 through 39 years (1990); percentage of the population aged 40 years or older (1990); male-to-female ratio (1990); percentage of the population non-Hispanic Black (1990); percentage of the population Hispanic (1990); percentage of the population urban (1990); birth rate per 1000 persons (1984); percentage of births to women younger than 20 years (1984); infant mortality rate per 1000 live births (1984); number of female heads of household per 1000 persons (1990); reported violent crime rate per 1000 persons (1985); percentage of the population living below the poverty level (1990); percentage of the civilian labor force unemployed (1990); percentage of adults with less than a ninth grade education (1990); divorce rate per 1000 persons (1984); local per capita expenditures for education (1992); local per capita expenditures for health and hospitals (1992); number of active, nonfederal physicians per 100 000 persons (1985); and number of hospital beds per 100 000 persons (1985). Since counties with high syphilis rates were concentrated in the South (defined by the US Bureau of the Census as Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Oklahoma, North Carolina, South Carolina, Tennessee, Texas, Virginia, West Virginia, and the District of Columbia), county location in the South was examined as an additional dichotomous variable.

Spearman's rank correlation coefficients were calculated to examine the association between the county sociodemographic variables and the county 10-year mean annual incidence of syphilis. This bivariate analysis was done for the following strata: all counties, non-Southern counties, Southern counties, Southern counties with a population of 250 000 or less, and Southern counties with a population of more than 250 000.

A multivariate linear regression analysis was performed, using a backward stepwise elimination process with a threshold P value of .001 for removal of variables from and retention of variables in the model. This high threshold ($P < .001$) was chosen to create a more parsimonious model. To reduce the variance of the outcome variable, it was transformed by raising it to the 0.3 power. This transformation value was selected by an iterative process that sought the transformation that most nearly resulted in a normal distribution of the residuals from the model.[13] The data set was randomly divided into two equal parts, and the transformation value was determined for one half of the data set and validated with the other half. The results are reported for the full data set. To reduce the variance in the model, we used the natural log of independent variables with a range of more than 100. If the range included 0, 1 was added to the variable value before taking the natural log. The regression was also done for the strata of non-Southern counties and Southern counties.

Results

From 1984 through 1993, a total of 355 783 cases of primary- or secondary-stage syphilis were reported from the 3085 US counties in the analysis. Of the 3085 counties, 896 (29%) had no cases during the 10-year period; 1084 (35%) had 1 to 9 cases; 737 (24%) had 10 to 99 cases; 306 (10%) had 100 to 999 cases; and 62 (2%) had 1000 or more cases. The 62 counties with 1000 or more cases accounted for 28% of the 1990 US population and for 65% of the syphilis cases. The 10-year mean annual incidence of syphilis per 100 000 persons ranged from 0 (896 counties) to 140. The mean of the 10-year mean rates was 7.0 and the median was 1.1. In 610 counties (20%), the 10-year mean annual incidence was higher than the 1990 *Healthy People 2000* goal[14] of 10 or fewer syphilis cases per 100 000 persons. Of these 610 counties, 550 (90%) were in the South (Figure 1). In the 2087 counties where 50% or less of the population was urban, the 10-year mean annual incidence was 5.9; in the 998 counties where more than 50% of the population was urban, the mean annual incidence was 9.4.

In the analysis including all counties, syphilis incidence was highly correlated ($r \geq .4$) with percentage of the population non-Hispanic Black, number of female heads of household per 1000 persons, rate of violent crime, and percentage of births to women younger than 20 years (Table 1). There was a moderate positive correlation ($.2 \leq r < .4$) with percentage of the population aged 15 to 39 years, percentage of the population urban, percentage of the population living below the poverty level, percentage unemployed, divorce rate, and percentage of the population Hispanic. Syphilis incidence was moderately negatively correlated ($-.4 > r \geq -.2$) with percentage of the population aged 40 years or older, male-to-female ratio, and local per capita expenditures for education, and poorly correlated ($r < .2$) with number of physicians per 100 000 persons, birth rate, percentage of adults with less than a ninth grade education, infant mortality rate, number of hospital beds per 100 000 persons, and percentage of the population aged 14 years or younger. In the bivariate correlation analyses stratified by region and population, percentage of the population non-Hispanic Black, number of female heads of household per 1000 persons, and violent crime rate (except in counties in the South with a population of less than 250 000) remained highly correlated with syphilis rates.

In the multivariate analysis for counties with no missing census data (n = 2935 [95.1%]), the full model R^2 was .71, that is, 71% of the variation in syphilis rates among counties was accounted for by these sociodemographic variables (Table 2). The factors most highly associated with syphilis were percentage of the population non-Hispanic Black, county location in the South, percentage of the population urban, percentage of the population Hispanic, and percentage of births to women younger than 20 years.

The association between 10-year mean annual incidence of syphilis and the three leading sociodemographic factors from the multivariate models (percentage of the population non-Hispanic Black, county location in the South, and percentage of the population urban) is shown in Figure 2. In each of the strata of percentage of the population non-Hispanic Black and percentage of the population urban, syphilis rates were higher in the Southern counties.

In separate analyses, the syphilis rates in the non-Southern counties were strongly associated with percentage non-Hispanic Black, percentage Hispanic, and percentage urban (Table 3), and the model R^2 was .53. In Southern counties, percentage non-Hispanic Black, percentage of adults with less than a ninth grade education (protective, negative association), and percentage Hispanic were the most highly associated

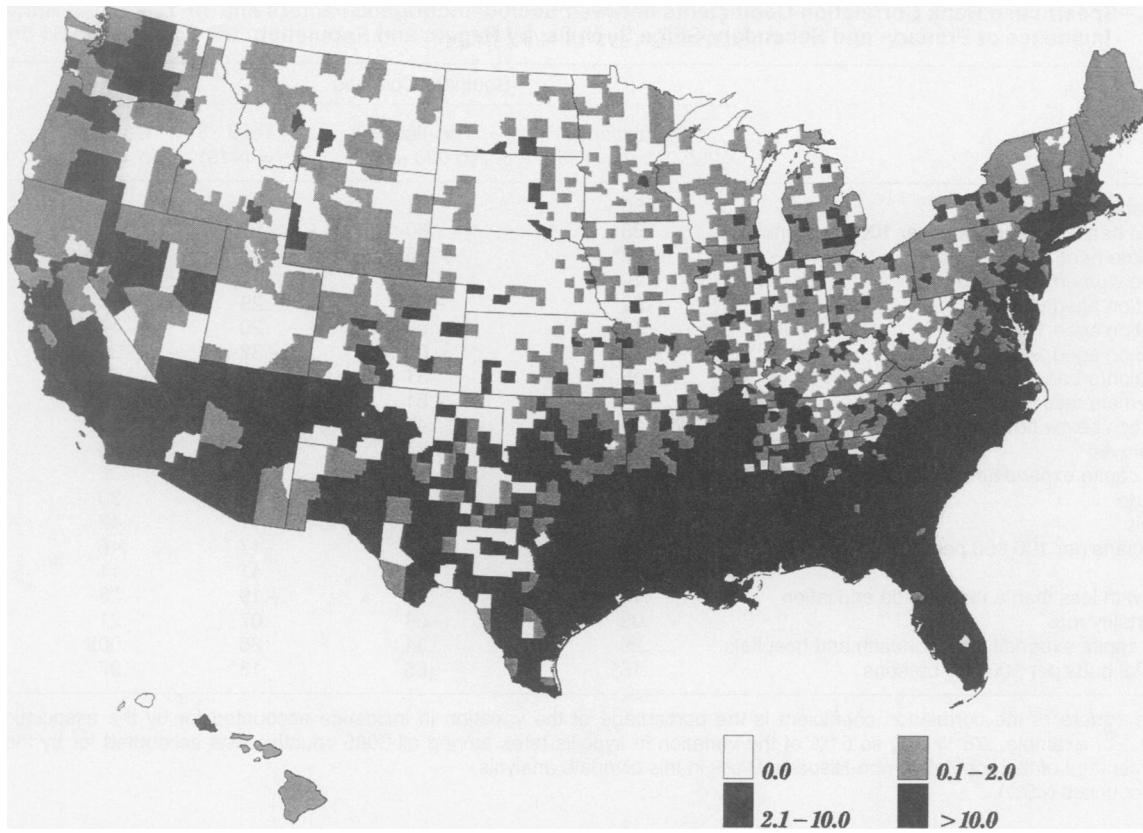


FIGURE 1—Mean annual incidence of primary- and secondary-stage syphilis per 100 000 persons, by county, United States, 1984 through 1993.

characteristics (Table 4), and the model R^2 was .65.

Discussion

In a multivariate model accounting for other sociodemographic factors, the county characteristics most strongly associated with primary- and secondary-stage syphilis rates were percentage non-Hispanic Black, county location in the South, percentage urban, percentage Hispanic, and percentage of births to women younger than 20 years. Sociodemographic characteristics accounted for 71% of the variation in syphilis rates among counties. The implications of these associations are considered below.

A county's percentage of non-Hispanic Blacks in the population was the characteristic most strongly associated with the county syphilis rate, even when accounting for the other factors in the multivariate analysis. Race is not thought to be a biological risk factor for syphilis. There are no studies showing different levels of susceptibility or infectiousness by racial or ethnic group, but, as is true for many other diseases, race/ethnicity is a marker of underlying determinants that account for the association. However, it is often very difficult to disentangle the interrelationships of race/ethnicity, other socioeconomic and demographic factors, and health status. Many studies of sexually transmitted diseases (STDs) (including CDC surveillance reports) include information on race, which is often easily available, but not on socioeconomic factors associated with both STDs and race. A stigma is attached to STDs and these reports can be misleading, resulting in a heightened sensitivity and distrust that make this a difficult issue to discuss.

In national surveillance data, the incidence of primary- and secondary-stage syphilis among Blacks has been approximately 60 times higher than that of Whites in recent years.[1] Some of the difference is due to Whites' greater use of private physicians, who may not report cases to the health department; Blacks are more likely to use public clinics, from which reporting is more nearly complete. In previous analyses done with the individual as the unit of analysis, other socioeconomic factors have accounted for some, but not all, of the rate difference between Blacks and Whites.[3] The association may be better explained by the dynamics of infectious disease epidemiology. Black and White Americans to a great extent compose separate sexual groups. In the University of Chicago National Health and Social Life Survey, 91% of all noncohabitational, heterosexual partnerships reported by Blacks and Whites were with partners of the same race.[15] Similar, although less strong, homophily (choice of partner with the same characteristics as oneself) was seen for education and religion, further reinforcing the separateness of sexual networks. Because the risk of acquiring infection depends upon infection prevalence among potential partners, STD epidemiology is nonlinear; factors associated with small differences in risk among individuals may generate large differences in risk between groups.[9] Even modest differences in the prevalence of risk factors between sexual networks may result in large differences in STD rates between these groups.

TABLE 1—Spearman's Rank Correlation Coefficients between Sociodemographic Factors and 10-Year Mean Annual Incidence of Primary- and Secondary-Stage Syphilis, by Region and Population: US Counties, 1984 through 1993

	Southern Counties			Non-Southern Counties (n = 1666)	Total US Counties (n = 3085)
	Population <250 000 (n = 1357)	Population ≥ 250 000 (n = 62)	Total (n = 1419)		
% Non-Hispanic Black	.80	.71	.80	.55	.78
No. female heads of household per 1000 persons	.66	.60	.67	.50	.61
Violent crime rate	.38[a]	.76	.42[a]	.54[a]	.50[a]
% Births to women aged <20 y	.19	.48	.15	.20	.45
% Population aged ≤14 y	.32	−.09	.29	−.08	.03
% Population aged 15–39 y	.17	.06	.20	.47	.36
% Population aged ≥40 y	−.32	.05	−.32	−.36	−.31
% Population urban	.26	.31	.30	.53	.33
Male-to-female ratio	−.32	−.31	−.32	−.09	−.31
% Population below poverty level	.19	.41	.15	−.09	.25
% Unemployed	.12	.35	.10	.23	.23
Local per capita expenditures for education	.12[a]	−.14	.11[a]	−.06[a]	−.21[a]
Divorce rate	−.05	.08	−.02	.30	.20
% Hispanic	.09	−.08	.11	.42	.20
No. physicians per 100 000 persons	.12	.50	.17	.46	.18
Birth rate	.41	.05	.41	.11	.17
% Adults with less than a ninth grade education	−.15	.23	−.19	−.23	.13
Infant mortality rate	.08	−.08	.07	.21	.12
Local per capita expenditures for health and hospitals	.25	.34	.26	.005	.08
No. hospital beds per 100 000 persons	.15[a]	.55	.18[a]	.07	.04[a]

Note. The square of the correlation coefficient is the percentage of the variation in incidence accounted for by the association with each variable. For example, $.78^2 = .61$, so 61% of the variation in syphilis rates among all 3085 counties was accounted for by the association with percentage of the population non-Hispanic Black in this bivariate analysis.
[a]Missing counties (<5%).

Some differences in STD risk factors by race and ethnicity have been identified. Fundamentally, STD rates depend upon transmissibility, rates of partner change, and duration of infectiousness.[16] In the National Health and Social Life Survey, having more than one sexual partner in the past 12 months was reported by 27% of Black, 20% of Hispanic, and 15% of White Americans,[15] although the difference between Blacks and Whites is much greater for men than for women.[17] Studies of other risky sexual behaviors and other national surveys have shown similar trends.[18–20] Access to medical care is also poorer for non-Whites. For example, whereas 83% of all adults had a specific source of primary care in 1993, only 79% of Blacks and 71% of Hispanics did.[21] There is also some evidence that Blacks wait longer than Whites to obtain medical care for STDs,[22,23] which would increase the duration of infectiousness. The crack cocaine epidemic and the accompanying exchange of sex for drugs was identified as a critical factor leading to the syphilis epidemic during the time of this study.[5,6] Crack cocaine use was more prevalent among Blacks[24] and may have been a factor in the higher syphilis rates among Blacks. Distrust of the public health system may further limit the effectiveness of control programs in Black communities by delaying patients' care seeking or decreasing the effectiveness of sex partner identification and treatment by public health workers.[25] These and other factors may each increase individual risk to some degree, resulting in a greatly increased disease burden in groups with a higher prevalence of these risk factors. While these data address racial/ethnic groups as a whole, syphilis remains an uncommon disease in all racial/ethnic groups, mainly affecting the most marginalized members of society.

In the national surveillance data, the incidence of primary- and secondary-stage syphilis was 5 to 10 times higher among Hispanics than among Whites during the time of this study.[1] The percentage of Hispanics in the population was associated with increased county syphilis rates in our ecological analysis, possibly for the same reasons that apply to the non-Hispanic Black population, that is, prevalence of some risky sexual behaviors,[15] poor access to medical care,[21] or use of crack cocaine.[24] There is more evidence of poorer access to medical care for Hispanics than there is of more risky sexual behavior.

County location in the South was also strongly associated with increased county syphilis rates, even accounting for percentage non-Hispanic Black and other factors in the multivariate model. There may have been regional differences in risk factors or in disease control program factors not accounted for by the other variables. Or, like the group effects of risk factors in populations, the concentration and contiguousness in the South of counties with characteristics associated with higher syphilis rates (such as percentage of non-Hispanic Blacks in the population) may generate higher rates in the region than individual county characteristics alone would predict. Because of importation of cases from neighboring counties, a county located in the South may have higher syphilis rates than if it were in another region surrounded by counties with low rates.

Urban residence has been shown to be associated with syphilis rates at the individual level in some studies.[26] Percentage of the population that was urban was strongly associated with county syphilis rates in this analysis, possibly because of the predominance in urban areas of the crack cocaine epidemic and of sexual risk behaviors,[20] including sex with anonymous partners (who would have been more difficult for public health workers to locate and treat). There may also be a surveillance artifact, with better access to medical care in urban areas[21] resulting in higher rates of appropriate diagnosis and reporting of syphilis cases. In the 1990s the syphilis epidemic reached rural areas in the South,[27] and in

this study syphilis rates were less strongly associated with the percentage of the population that was urban in Southern counties than in counties outside the South.

County syphilis rates were associated with the percentage of births to women younger than 20 years but not with overall birth rates, so this factor may not simply be a marker of unprotected heterosexual intercourse. At the individual level, the syphilis rate in women is highest in 20- to 24-year-olds, not in teenagers.[1] The percentage of births to women younger than 20 years may therefore be a marker for other social conditions that lead to higher syphilis rates, such as high-risk sexual behavior or delay in seeking care for STDs.

The numbers of physicians and hospital beds per 100 000 persons, infant mortality rate, and local per capita expenditures for health and hospitals may not be good markers of access to care for STDs, but it is important to note that these factors were not protective against high syphilis rates in the multivariate analysis. This does not obviate the need to improve access to care as an important part of an STD control program, but it does suggest that the presence of health care facilities and workers alone does not control syphilis. The number of physicians per capita was *positively* associated with syphilis rates in the bivariate analysis and remained positively associated in the multivariate model. This may be a surveillance artifact, in that syphilis must be diagnosed by a health care provider to be reported.

Imbalance in community sex ratios is thought to lead to sexual behaviors that increase STD incidence.[28] In the United States, the low male-to-female ratio in many Black communities[29] may allow men to have multiple female sex partners and increase STD rates. Indeed, in the bivariate analysis, male-to-female ratio was moderately negatively associated with county syphilis rates. However, this factor was eliminated from the multivariate model when other factors were taken into account.

There are several limitations to this study. First, it was limited to the available county-level data. No reliable national data on sexual behavior or crack cocaine use, both of which could have been important analysis variables, are available at the county level. Also, while STD control programs are usually organized at the county level (as are the available national data), sexual networks may not coincide with these political boundaries. Analyses using smaller areas, such as zip codes or census tracts, would provide additional information. For example, in a study of syphilis in

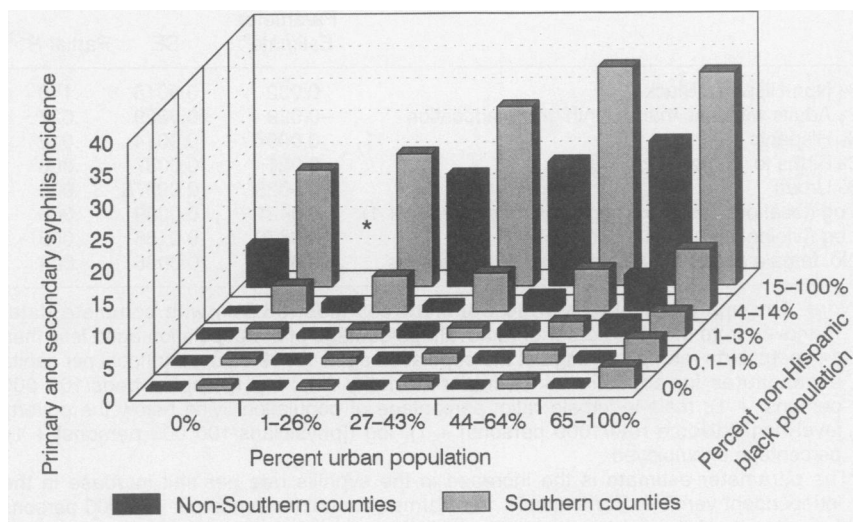
TABLE 2—Backward Elimination Multivariate Regression of 20 Sociodemographic Factors (Independent Variables) and 10-Year Mean Annual Incidence of Primary- and Secondary-Stage Syphilis (Dependent Variable): US Counties, 1984 through 1993

	Parameter Estimate[a]	SE	Partial R^2 [b]
% Non-Hispanic Black	0.0350	0.0012	.0899
South region (vs. non-South)	0.4332	0.0284	.0227
% Urban	0.0050	0.0004	.0130
% Hispanic	0.0101	0.0010	.0102
% Births to women aged <20 y	0.0202	0.0025	.0062
% Adults with less than a ninth grade education	–0.0138	0.0020	.0049
Log ([physicians/100 000 persons] + 1)	0.0596	0.0101	.0034
Log ([violent crime/1000 persons] + 1)	0.0506	0.0089	.0032
No. female heads of household per 1000 persons	0.0133	0.0024	.0030
% Unemployed	–0.0179	0.0044	.0016

Note. Removed from model (in order of removal): male-to-female ratio; log ([hospital beds/100 000 persons] + 1); log ([divorce rate/1000 persons] + 1); percentage of the population aged 15–39 years; log (local per capita expenditures for education + 1); infant mortality rate per 1000 live births; percentage of population living below the poverty level; log (local per capita expenditures for health and hospitals + 1); percentage of the population aged less than 14 years; and birth rate.

[a]The parameter estimate is the increase in the syphilis rate per unit increase in the independent variable. For example, transformed annual incidence per 100 000 persons (raised to the 0.3 power) increases by .0350 for every 1% increase in the percentage of the county's population that is non-Hispanic Black.

[b]The partial R^2 is the increase in the model R^2 when the variable is added as the last variable in the model. For example, an additional 8.99% of the variation in syphilis rates among counties was accounted for by the addition to the model of percentage non-Hispanic Black population.



Note. The height of the bar represents the average of the syphilis rates of the counties in each stratum.
*No counties in this stratum.

FIGURE 2—Association of 10-year mean annual incidence of primary- and secondary-stage syphilis per 100 000 persons in US counties, 1984 through 1993, with percentage non-Hispanic Black population, percentage of the population that was urban, and region.

TABLE 3—Backward Elimination Multivariate Regression of 19 Sociodemographic Factors (Independent Variables) and 10-Year Mean Annual Incidence of Primary- and Secondary-Stage Syphilis (Dependent Variable): Non-Southern US Counties, 1984 through 1993

	Parameter Estimate[a]	SE	Partial R^{2}[b]
% Non-Hispanic Black	0.054	0.0030	.093
% Hispanic	0.015	0.0014	.031
% Urban	0.0042	0.00056	.018
Log ([violent crime/1000 persons] + 1)	0.049	0.0095	.008
% Births to women aged <20 y	0.015	0.0029	.008
Log ([physicians/100 000 persons] + 1)	0.064	0.0136	.007
% Population aged 15–39 y	1.19	0.282	.005

Note. The analysis includes 1569 non-Southern counties (94.2%) with complete data. Removed from model (in order of removal): log ([divorce rate/1000 persons] + 1); log (local per capita expenditures for health and hospitals + 1); percentage unemployed; log (local per capita expenditures for education + 1); percentage of the population aged less than 14 years; birth rate; log ([hospital beds/100 000 persons] + 1); infant mortality rate; male-to-female ratio; percentage of population living below the poverty level; percentage of adults with less than a ninth grade education; and number of female heads of household per 1000 persons.

[a]The parameter estimate is the increase in the syphilis rate per unit increase in the independent variable. For example, transformed annual incidence per 100 000 persons (raised to the 0.3 power) increases by 0.054 for every 1% increase in the percentage of the county's population that is non-Hispanic Black.

[b]The partial R^{2} is the increase in the model R^{2} when the variable is added as the last variable in the model. For example, an additional 9.3% of the variation in syphilis rates among counties was accounted for by the addition to the model of percentage non-Hispanic Black population.

TABLE 4—Backward Elimination Multivariate Regression of 19 Sociodemographic Factors (Independent Variables) and 10-Year Mean Annual Incidence of Primary- and Secondary-Stage Syphilis (Dependent Variable): Southern US Counties, 1984 through 1993

	Parameter Estimate[a]	SE	Partial R^{2}[b]
% Non-Hispanic Black	0.032	0.0016	.110
% Adults with less than a ninth grade education	–0.029	0.0029	.027
% Hispanic	0.0094	0.0014	.012
% Births to women aged <20 y	0.021	0.0037	.008
% Urban	0.0038	0.00075	.006
Log (local per capita expenditures for education + 1)	0.0421	0.0089	.006
Log ([violent crime/1000 persons] + 1)	0.0670	0.0156	.005
No. female heads of household per 1000 persons	0.0145	0.0036	.004

Note. The analysis includes 1366 Southern counties (96.3%) with complete data. Removed from model (in order of removal): percentage of the population aged less than 14 years; birth rate; percentage of the population aged 15–39 years; log (local per capita expenditures for education + 1); infant mortality rate; log ([hospital beds/100 000 persons] + 1); male-to-female ratio; percentage of population living below the poverty level; log ([divorce rate/1000 persons] + 1); log ([physicians/100 000 persons] + 1); percentage unemployed.

[a]The parameter estimate is the increase in the syphilis rate per unit increase in the independent variable. For example, transformed annual incidence per 100 000 persons (raised to the 0.3 power) increases by 0.032 for every 1% increase in the percentage of the county's population that is non-Hispanic Black.

[b]The partial R^{2} is the increase in the model R^{2} when the variable is added as the last variable in the model. For example, an additional 11.0% of the variation in syphilis rates among counties was accounted for by the addition to the model of percentage non-Hispanic Black population.

Philadelphia using income data from the zip code of residence of syphilis patients, the ecological correlation between poverty and syphilis was strong even after accounting for race.[30] Another potential limitation is that STDs are underreported,[31] although in one study in rural North Carolina, all positive syphilis test results were reported.[32] Finally, some of the risk characteristics identified are intercorrelated; however, the association between syphilis and the factors most highly correlated with syphilis was stable with variations in the model.

Further investigation of the significance of these factors and the mechanisms by which they account for variation in syphilis rates is needed. We are studying 12 counties that were outliers from the multivariate model; that is, they had rates higher or lower than would be expected on the basis of their sociodemographic characteristics. This research is focused on the factors identified in this study and on other potential determinants for which national county-level data were not available, such as local syphilis prevention program attributes (including numbers of public STD clinicians, outcomes of sex partner notification and treatment efforts, and serological screening policies), transactional sex, illicit drug use, and other social, legal, economic, and political factors.

In assessing syphilis prevention programs, it is important to note that in this analysis sociodemographic characteristics accounted for most of the variation in syphilis rates among counties. While syphilis control programs have substantially reduced the burden of syphilis in industrialized countries, and the effectiveness of local syphilis prevention programs may be correlated with the factors identified in this study, it appears that differences between prevention programs may have accounted for only a small proportion of the variation in syphilis rates among counties. Our study of 12 counties with rates higher or lower than expected on the basis of their sociodemographic characteristics may identify counties in which the effectiveness of prevention programs has been an important determinant of syphilis rates and may identify the key attributes of effective programs. Since syphilis incidence is a marker of persons and populations at high risk for human immunodeficiency virus (HIV) infection, the epidemiology of syphilis may provide information about the HIV epidemic that would otherwise be recognized only after a long latent period and that could be used to target and evaluate prevention programs.[33] Unlike HIV infection, syphilis is often diagnosed and reported in its early stages, and county-level data on syphilis, but not HIV infection or other STDs, are consistently reported to the CDC.

Counties with high syphilis rates are also likely to be overburdened by other infectious and chronic diseases and by social problems including single-parent households, violent crime, and poverty.

Addressing these larger community issues may be necessary to improve the overall health status of these populations. □

Acknowledgments

The authors thank Sharon Hixon, Keith Humphrey, and Ray Ransom for their assistance in graphics production.

References

1. *Sexually Transmitted Disease Surveillance, 1994.* Atlanta, Ga: Centers for Disease Control and Prevention, Division of STD Prevention; September 1995.
2. Nakashima AK, Rolfs RT, Flock ML, Kilmarx P, Greenspan JR. Epidemiology of syphilis in the United States, 1941–1993. *Sex Transm Dis.* 1996;23:16–23.
3. Hahn RA, Magder LS, Aral SO, Johnson RE, Larsen SA. Race and the prevalence of syphilis seroreactivity in the United States population: a national sero-epidemiologic study. *Am J Public Health.* 1989;79:467–470.
4. Siegal D, Larsen SA, Golden E, Morse S, Fullilove MT, Washington AE. Prevalence, incidence and correlates of syphilis seroreactivity in multiethnic San Francisco neighborhoods. *Ann Epidemiol.* 1994;4:460–465.
5. Rolfs RT, Goldberg M, Sharrar RG. Risk factors for syphilis: cocaine and prostitution. *Am J Public Health.* 1990;80:853–857.
6. Farley TA, Hadler JL, Gunn RA. The syphilis epidemic in Connecticut: relationship to drug use and prostitution. *Sex Transm Dis.* 1990;17:163–168.
7. Susser M. The logic in ecological, I: the logic of analysis. *Am J Public Health.* 1994; 84:825–829.
8. Susser M. The logic in ecological, II: the logic of design. *Am J Public Health.* 1994;84: 830–835.
9. Koopman JS, Longini IM Jr. The ecological effects of individual exposures and nonlinear disease dynamics in populations. *Am J Public Health.* 1994;84:836–842.
10. Centers for Disease Control. Case definitions for public health surveillance. *MMWR Morb Mortal Wkly Rep.* 1990;39(RR-13):1–43.
11. *County and City Data Book, 1988.* Washington, DC: US Dept of Commerce, Bureau of the Census; 1988.
12. *Census of Population and Housing, 1990.* Washington, DC: US Dept of Commerce, Bureau of the Census; 1991.
13. Neter J, Kutner MH, Nachtsheim CL, Wasserman W. *Applied Linear Regression Models.* 3rd ed. Chicago, Ill: Irwin; 1996:132–134.
14. *Healthy People 2000: National Health Promotion and Disease Prevention Objectives.* Washington, DC: US Dept of Health and Human Services; 1991. DHHS publication PHS 91-50212.
15. Laumann EO, Gagnon JH, Michael RT, Michaels S. The social organization of sexuality: sexual practices in the United States. Chicago, Ill: The University of Chicago Press; 1994.
16. Brunham RC, Plummer FA. A general model of sexually transmitted disease epidemiology and its implications for control. *Med Clin North Am.* 1990;74:1339–1352.
17. Seidman SN, Aral S. Subpopulation differentials in STD transmission. *Am J Public Health.* 1992;82:1297. Letter.
18. Centers for Disease Control and Prevention. Youth risk behavior surveillance—United States, 1993. *MMWR CDC Surveill Summ.* 1995;44(No. SS-1):1–56.
19. Choi KH, Catania JA, Dolcini MM. Extramarital sex and HIV risk behavior among US adults: results from the National AIDS Behavioral Survey. *Am J Public Health.* 1994;84:2003–2007.
20. Davis JA, Smith TW. *General Social Surveys, 1972–1993: Cumulative Codebook.* Chicago, Ill: National Opinion Research Center; 1993.
21. Centers for Disease Control and Prevention. 1992 National Health Interview Survey. In: *Healthy People 2000 Review.* Washington, DC: US Dept of Health and Human Services; 1994.
22. Darrow WW. Venereal infections in three ethnic groups in Sacramento. *Am J Public Health.* 1976;66:446–450.
23. Irwin DE, Thomas JC, Leone PA, et al. Sexually transmitted disease patients' self-treatment practices prior to seeking medical care. In: *Abstracts of the 35th Interscience Conference on Antimicrobial Agents and Chemotherapy (San Francisco).* Washington, DC: American Society for Microbiology; 1995:301. Abstract K74.
24. *National Household Survey on Drug Abuse: Population Estimates, 1993.* Rockville, Md: Substance Abuse and Mental Health Services Administration; October 1994.
25. Thomas SB, Quinn SC. The Tuskegee syphilis study, 1932 to 1972: implications for HIV education and AIDS risk education programs in the black community. *Am J Public Health.* 1991;81:1498–1505.
26. Rothenberg R. The geography of syphilis: a demonstration of epidemiologic diversity. In: Morisset R, Kourstak E, eds. *Advances in Sexually Transmitted Diseases, Diagnosis and Treatment.* Utrecht, the Netherlands: VNU Sciences Press; 1986:125–133.
27. Thomas JC, Kulik AL, Schoenbach VJ. Syphilis in the South: rural rates surpass urban rates in North Carolina. *Am J Public Health.* 1995;85:1119–1122.
28. Aral SO. The social context of syphilis persistence in the southeastern United States. *Sex Transm Dis.* 1996;23:9–15.
29. Tucker MB, Taylor RJ. Demographic correlates of relationship status among Black Americans. *J Marriage Fam.* 1989;51:655–665.
30. Toomey KE, Moran JS, Rafferty MP, Beckett GA. Epidemiological considerations of sexually transmitted diseases in underserved populations. *Infect Dis Clin North Am.* 1993;7:739–752.
31. Rothenberg R, Bross DC, Vernon TM. Reporting of gonorrhea by private physicians: a behavioral study. *Am J Public Health.* 1980;70:983–986.
32. Smucker DR, Thomas JC. Evidence of thorough reporting of sexually transmitted diseases in a southern rural county. *Sex Transm Dis.* 1995;22:149–154.
33. Otten MW, Zaidi AA, Peterman TA, Rolfs RT, Witte JJ. High rate of HIV seroconversion among patients attending sexually transmitted disease clinics. *AIDS.* 1994;8:549–553.